

WHAT IS CLAIMED IS:

1. A method of normalizing a discourse representation structure (DRS) that includes boxes with box elements and box element arguments having markers, box identifiers and data values, the method comprising:
 - normalizing a form in which boxes and box elements are represented in the DRS;
 - sorting the boxes and box elements, based on the box and box element normal form and regardless of the markers, to obtain a preliminary ordering;
 - normalizing a form in which markers are represented in the DRS; and
 - sorting the boxes and box elements based on the preliminary ordering and based on the marker normal form, to achieve a normal form for the DRS.
2. The method of claim 1 wherein normalizing the form in which boxes are represented comprises:
 - deleting unused boxes from the DRS.
3. The method of claim 2 wherein normalizing the form in which box elements are represented comprises:
 - deleting unused box elements from the DRS.
4. The method of claim 3 wherein normalizing the form in which boxes are represented comprises:
 - re-numbering the boxes with consecutive indices.

5. The method of claim 4 wherein normalizing the form in which box elements are represented comprises:
re-numbering the box elements with consecutive identifiers.

6. The method of claim 5 wherein sorting the boxes and box elements, based on the box and box element normal forms and regardless of the markers, comprises:

lexicographically ordering the boxes and box elements based on the indices and identifiers to obtain lexicographically ordered boxes and box elements.

7. The method of claim 6 wherein sorting the boxes and box elements, based on the box and box element normal forms and regardless of the markers, comprises:

updating the boxes to refer to re-numbered box elements.

8. The method of claim 7 wherein sorting the boxes and box elements, based on the box and box element normal forms and regardless of the markers, comprises:

updating the box elements to refer to re-numbered boxes.

9. The method of claim 8 wherein normalizing the form in which markers are represented in the DRS comprises:

generating a mapping between each marker and a list identifying a box and box element containing the marker.

10. The method of claim 9 wherein normalizing the form in which markers are represented in the DRS comprises:

generating an inverse mapping between the list identifying a box and box element containing a marker and each marker.

11. The method of claim 10 wherein normalizing the form in which markers are represented in the DRS comprises:

re-numbering the markers with consecutive marker values.

12. The method of claim 11 wherein normalizing the form in which markers are represented in the DRS comprises:

updating the box elements to refer to the re-numbered markers.

13. The method of claim 12 wherein sorting the boxes and box elements based on the preliminary ordering and based on the marker normal forms comprises:

sorting the lexicographically ordered boxes and box elements based on the re-numbered markers to obtain a normalized DRS.

14. The method of claim 13 and further comprising:
generating a string representative of the
normalized DRS.
15. A discourse representation data structure (DRS)
representative of a discourse input, the DRS
comprising:
an array of boxes, each box including a set of
box elements with associated arguments, the
box elements and associated arguments
including a semantic representation of
semantic content of the discourse input.
16. The DRS of claim 15 and further comprising:
a string representative of the DRS.
17. The DRS of claim 15 and further comprising:
an integer number of boxes in the DRS indicative
of a length of the array of boxes.
18. The DRS of claim 17 wherein the set of box
elements is embodied as a vector of box
elements and further comprising:
an integer number of box elements in the DRS
indicative of a length of the vector of box
elements.

19. The DRS of claim 18 and further comprising:
a vector map that includes a vector of lists of
pairs of integers in the DRS; and
an integer number of markers in the DRS,
indicative of a length of the vector map.
20. The DRS of claim 15 wherein each box in the
array of boxes comprises:
a DRS pointer field that includes a pointer to
the DRS containing the box.
21. The DRS of claim 20 wherein each box in the
array of boxes comprises:
an integer number of elements in the box.
22. The DRS of claim 21 wherein each box in the
array of boxes comprises:
a vector of element indices having a length
given by the integer number of elements in
the box.
23. The DRS of claim 15 wherein each box element in
the set of box elements comprises:
a kind field indicative of a kind of box
element; and
a semantic kind field indicative of a semantic
kind of the box element.

24. The DRS of claim 23 wherein each box element further comprises:

an integer number of box element arguments in the box element; and

a vector of box element arguments in the box element, having a length indicated by the integer number of box element arguments.

25. The DRS of claim 24 wherein each box element further comprises:

a string representative of the box element.

26. The DRS of claim 24 wherein each box element further comprises:

a semantic node value indicative of a semantic node in an external semantic domain corresponding to the box element.

27. The DRS of claim 15 wherein each box element argument comprises:

an argument kind field indicative of a kind of the box element argument.

28. The DRS of claim 27 wherein each box element argument comprises:

an argument identification field including an identifier of the box element argument.

29. The DRS of claim 27 wherein each box element argument comprises:

